Computers in Biology and Medicine 71 (2016) 24–34

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Phonocardiogram signal compression using sound repetition and vector quantization

Hong Tang [a,*], Jinhui Zhang [a], Jian Sun [a], Tianshuang Qiu [a], Yongwan Park [b]

[a] Department of Biomedical Engineering, Dalian University of Technology, Dalian, PR China
[b] Department of Information and Communication Engineering, Yeungnam University, Dae-dong, Gyeongsan-si, Gyeongsangbuk-do 712749, Republic of Korea

ARTICLE INFO

Article history:
Received 23 July 2015
Accepted 14 January 2016

Keywords:
Phonocardiogram signal
Signal compression
Sound repetition
Time–frequency decomposition
Vector quantization

ABSTRACT

Background: A phonocardiogram (PCG) signal can be recorded for long-term heart monitoring. A huge amount of data is produced if the time of a recording is as long as days or weeks. It is necessary to compress the PCG signal to reduce storage space in a record and play system. In another situation, the PCG signal is transmitted to a remote health care center for automatic analysis in telemedicine. Compression of the PCG signal in that situation is necessary as a means for reducing the amount of data to be transmitted. Since heart beats are of a cyclical nature, compression can make use of the similarities in adjacent cycles by eliminating repetitive elements as redundant. This study proposes a new compression method that takes advantage of these repetitions.
Methods: Data compression proceeds in two stages, a training stage followed by the compression as such. In the training stage, a section of the PCG signal is selected and its sounds and murmurs (if any) decomposed into time–frequency components. Basic components are extracted from these by clustering and collected to form a dictionary that allows the generative reconstruction and retrieval of any heart sound or murmur. In the compression stage, the heart sounds and murmurs are reconstructed from the basic components stored in the dictionary. Compression is made possible because only the times of occurrence and the dictionary indices of the basic components need to be stored, which greatly reduces the number of bits required to represent heart sounds and murmurs. The residual that cannot be reconstructed in this manner appears as a random sequence and is further compressed by vector quantization. What we propose are quick search parameters for this vector quantization.
Results: For normal PCG signals the compression ratio ranges from 20 to 149, for signals with median murmurs it ranges from 14 to 35, and for those with heavy murmurs, from 8 to 20, subject to a degree of distortion of $\sim$5% (in percent root-mean-square difference) and a sampling frequency of 4 kHz.
Discussion: We discuss the selection of the training signal and the contribution of vector quantization. Performance comparisons between the method proposed in this study and existing methods are conducted by computer simulations.
Conclusions: When recording and compressing cyclical sounds, any repetitive components can be removed as redundant. The redundancies in the residual can be reduced by vector quantization. The method proposed in this study achieves a better performance than existing methods.

1. Background

Heart sounds are generated by the interactions between heart chambers, valves and great vessels and the blood flowing through them. Mechanical vibrations reflect the turbulence that occurs when heart valves close. Traditionally, a stethoscope is used in cardiac auscultation to listen to these sounds that provide important acoustic information regarding the condition of the heart. In home health monitoring applications, a phonocardiogram (PCG) signal may be continuously recorded for hours, days or even weeks to catch events related to heart hemodynamics. If a single channel PCG signal is collected for one day with a sampling frequency of 4 kHz and a digitization depth of 16 bits, the data storage space required is 24*60*60*4000*16 bits $\approx$ 0.69 GBytes. More storage space is needed if multi-channel signals are collected over several days. To save storage space in a record and play system, it is necessary to compress the data. Furthermore, the rapidly developing field of telemedicine has made it possible to transmit a PCG record to a health center for automatic remote analysis. Let us consider a scenario, as shown in Fig. 1. A person with a smart

* Corresponding author. Tel.: +86 411 84706009x3013.
E-mail address: tanghong@dlut.edu.cn (H. Tang).

http://dx.doi.org/10.1016/j.compbiomed.2016.01.017
0010-4825/

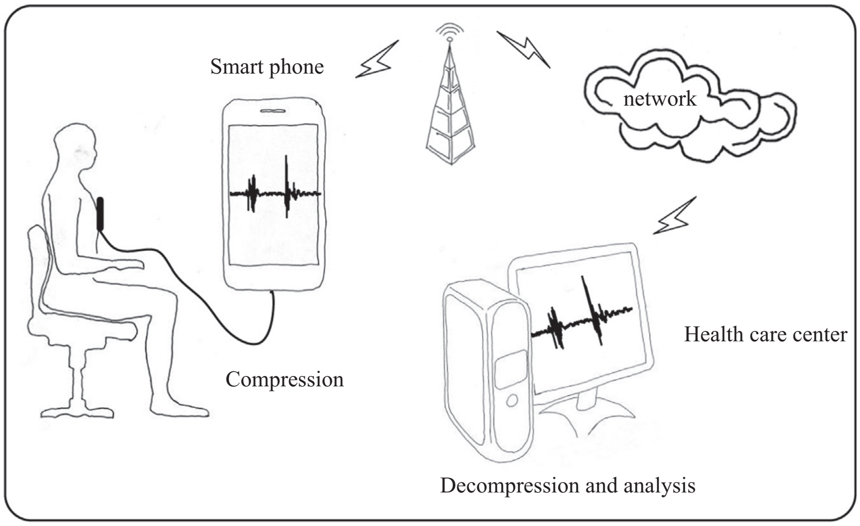

Fig. 1. Illustration of the remote analysis of a PCG signal in a mobile medicine application.

phone can access a remote health care service anywhere and anytime. The PCG signal (or other physiological signal) can be continuously recorded by attaching a PCG sensor at the required position. The data are compressed by the embedded algorithm in the smart phone. Remote analysis benefits from improved transmission efficiency via the compression technique. The PCG signal is decompressed at the health care center without any loss of physiological information.

To achieve the highest possible compression ratio with the lowest possible distortion, many researchers have developed compression methods for PCG data. The first reported method to compress a PCG signal used a wavelet transform (WT) [1]. The PCG signal was divided into non-overlapping blocks, and the wavelet transform was applied to each block independently. A hard threshold was applied to select the prominent wavelet coefficients, and the PCG signal was reconstructed from the reduced number of coefficients and their positions in the wavelet domain. To obtain the optimal wavelet, block length, and WT decomposition level for the compression, the authors used a genetic algorithm [2]. The compression was later implemented in an FPGA-based system [3]. It can be concluded that the smaller the number of selected coefficients, the higher the compression ratio of the method will be. However, the wavelet coefficients are not sparse enough, and to achieve an acceptably low distortion, a great number of coefficients must be selected. The computer experiments showed that the compression ratio was roughly 6 for a normal PCG signal at a sampling frequency of 8 kHz subject to a distortion of 6.7% in terms of the percent root-mean-square difference. Based on this reasoning, a transform was desired that would decompose a PCG signal into sparse components. Modified WT-based methods appeared in the following years and exhibited a similar performance [4–9]. Another hard threshold method based on adaptive Fourier decomposition was proposed in [10]. In addition to the lossy compression methods mentioned above, lossless compression in [11] using the Lempel–ziv–storer–szymanski technique was applied in a remote heart sound monitoring system. However, the compression ratios of lossless methods are generally lower than those of lossy methods. Consequently, more data needs to be transmitted over the network, while at the same time it is necessary to reduce the amount of data transmitted without the loss of physiological information.

The objective of this study was to achieve a higher compression ratio and lower degree of distortion when storing or transmitting heart sound signals. This was accomplished by using a previously undocumented approach whereby those sounds or murmurs are discarded as redundant that recur from one cardiac cycle to another. The heart sounds or murmurs are decomposed into their time–frequency components and the basic components are then extracted by clustering to produce a dictionary. The heart sounds or murmurs can then be reconstructed from their basic components, where for each component only the dictionary index and the time of occurrence need to be recorded. This strategy greatly reduces the number of bits required to represent heart sounds or murmurs. The residual, that part of the signal that cannot be reconstructed by the basic components, is further compressed by vector quantization, which results in a better fidelity.

2. Methods

2.1. Dictionary formulations

2.1.1. Time–frequency decomposition

Heart sounds and/or murmurs are fast, time-varying waveforms that can be decomposed into components in a joint time–frequency domain. Decomposition can be achieved through various methods, such as the discrete wavelet-based method [1–9], the matching pursuit method [12,13], or adaptive time–frequency decomposition [14,15]. The discrete wavelet-based method and the matching pursuit method both produce non-sparse components, typically many more than are produced by adaptive time–frequency decomposition. So in this study we adopted the adaptive time–frequency decomposition since it produces a very sparse representation of heart sounds and murmurs. It is written as

$$x(t) = \sum_{i=1}^{M} a_i e^{(t-t_i)^2/(2\sigma_i^2)} \cos(2\pi f_i t + \beta_i), \quad (1)$$

where $x(t)$ is a heart sound signal, which is considered to be the sum of M components. Each component is characterized by five parameters, namely amplitude, a_i; occurrence time, t_i; frequency, f_i; time-support, σ_i; and phase, β_i. The i-th component can be determined by a parameter vector $[a_i, t_i, f_i, \sigma_i, \beta_i]$. The presentation of (1) is similar to the short time Fourier transform (STFT)

$$H(t,f) = \int h(\tau) w(t-\tau) e^{-2\pi f \tau} d\tau, \quad (2)$$

where $h(t)$ is the signal to be analyzed. The sliding Gaussian window, $w(t)$, is centered at t_i and covers the heart sounds. Its